```
          |:|:    |  ||: : :||:   | |                  | ::|  :|  :|
Db    356 EVSEEDVHDMPYLKAVVLEGLRKHPPAHMLLPHKAAEDMDVGGYLIPKGTIVNFMVAEMG 415

Qy     91 SQ------------------DGAGTD 98
                              || | |
Db    416 RDEKEWEKPMEFMPERFLPGGDGEGVD 442


RESULT 14
O51827
ID   O51827      PRELIMINARY;      PRT;  2458 AA.
AC   O51827;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Polyketide synthase type I.
GN   PLTB.
OS   Pseudomonas fluorescens.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=294;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PF-5;
RX   MEDLINE=98094250; PubMed=9434161;
RA   Nowak-Thompson B., Gould S.J., Loper J.E.;
RT   "Identification and sequence analysis of the genes encoding a
RT   polyketide synthase required for pyoluteorin biosynthesis in
RT   Pseudomonas fluorescens Pf-5.";
RL   Gene 204:17-24(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PF-5;
RX   MEDLINE=99194726; PubMed=10094695;
RA   Nowak-Thompson B., Chaney N., Wing J.S., Gould S.J., Loper J.E.;
RT   "Characterization of the pyoluteorin biosynthetic gene cluster of
RT   Pseudomonas fluorescens Pf-5.";
RL   J. Bacteriol. 181:2166-2174(1999).
DR   EMBL; AF081920; AAC38075.1; -.
DR   InterPro; IPR001227; Ac_transferase.
DR   InterPro; IPR000794; Ketoacyl-synt.
DR   InterPro; IPR003880; Ppantne_attach.
DR   InterPro; IPR002155; Thiolase..
DR   Pfam; PF00698; Acyl_transf; 1.
DR   Pfam; PF00109; ketoacyl-synt; 2.
DR   Pfam; PF02801; ketoacyl-synt_C; 2.
DR   Pfam; PF00550; pp-binding; 2.
DR   PROSITE; PS50075; ACP_DOMAIN; 2.
DR   PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
DR   PROSITE; PS00098; THIOLASE_1; UNKNOWN_1.
KW   Phosphopantetheine; Transferase.
SQ   SEQUENCE   2458 AA;  262676 MW;  AE756080AE1A5FB1 CRC64;

  Query Match             12.8%;  Score 69.5;  DB 2;  Length 2458;
  Best Local Similarity   32.1%;  Pred. No. 2.4e+02;
  Matches   17;  Conservative   14;  Mismatches   19;  Indels    3;  Gaps    2;

Qy     44 ELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGSQDGAG 96
          :|   |  :  : :|| || | ::    ||||: ::::|: : :||| |   |
Db   2239 QLLHLTQSEQELEQDA-AARTAMQG--SGLQPLQRSQIVQAIARVLGGQGQCG 2288


RESULT 15
Q9CFY4
ID   Q9CFY4      PRELIMINARY;      PRT;   540 AA.
AC   Q9CFY4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Fibronectin-binding protein.
GN   YNGB OR LL1327.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Streptococcaceae; Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IL1403;
RX   MEDLINE=21235186; PubMed=11337471;
RA   Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA   Weissenbach J., Ehrlich S.D., Sorokin A.;
RT   "The complete genome sequence of the lactic acid bacterium Lactococcus
RT   lactis ssp. lactis IL1403.";
RL   Genome Res. 11:731-753(2001).
DR   EMBL; AE006365; AAK05425.1; -.
KW   Complete proteome.
SQ   SEQUENCE   540 AA;  61285 MW;  4DAC3CCCBCF72867 CRC64;

  Query Match             12.7%;  Score 69;  DB 16;  Length 540;
  Best Local Similarity   30.6%;  Pred. No. 41;
  Matches   22;  Conservative    9;  Mismatches   37;  Indels    4;  Gaps    2;

Qy     25 DNDEFFM---DFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMTELKSCIDGLQP-MHKAE 80
          || | |    : | | |   | :    ||  |  ||  : |:    | :  |   |: :
Db    294 DNAEIFRQKGELLNTFLNQVPNDKTSVTLENYYTNEPIEIALNPALSPVQNAQRYFHRYQ 353

Qy     81 LVKLLVQVLGSQ 92
          :|  |: || |
Db    354 KLKQAVKFLGEQ 365


Search completed: May 14, 2003, 10:52:40
Job time : 37 secs
```

```
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00300; SRP54; UNKNOWN_1.
SQ   SEQUENCE   500 AA;  55492 MW;  C957B6F76C6A224B CRC64;

  Query Match             13.4%;  Score 72.5;  DB 11;  Length 500;
  Best Local Similarity   32.4%;  Pred. No. 15;
  Matches   36;  Conservative   14;  Mismatches   42;  Indels   19;  Gaps    6;

Qy       2 KGSRALLLVALTLFCICRMATGEDNDEF----FMDFLQTLL-VGTPEELYEGTLGKYNVN 56
           ||  ||  || |   |  : ||| | |        |:  || :|  | | :     ||
Db     255 KGGGALSAVAATKSPIIFIGTGEHIDNFEPFKTQPFISKLLGMGDIEGLID------KVN 308

Qy      57 E----DAKAAMTELKSCIDGLQPMHK--AELVKL--LVQVLGSQDGAGTDY 99
           |    | || : : |     |: |::    ::|:    |:||   | |||:
Db     309 ELKLDDNKALIEKWKHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDF 359


RESULT 11
Q8QXL2
ID   Q8QXL2      PRELIMINARY;      PRT;  3071 AA.
AC   Q8QXL2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Polyprotein.
OS   Sorghum mosaic virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC   Potyvirus.
OX   NCBI_TaxID=32619;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=XIAOSHAN;
RA   Chen J., Chen J.P., Adams M.J.;
RT   "Characterisation of potyviruses from sugarcane and maize in China.";
RL   Arch. Virol. 147:0-0(2002).
DR   EMBL; AJ310197; CAC84437.1; -.
FT   CHAIN         2    233       P1 PROTEIN.
FT   CHAIN       234    693       HC-PRO PROTEIN.
FT   CHAIN       694   1040       P3 PROTEIN.
FT   CHAIN      1041   1107       6K1 PROTEIN.
FT   CHAIN      1108   1745       CI PROTEIN.
FT   CHAIN      1746   1798       6K2 PROTEIN.
FT   CHAIN      1799   1987       NIA-VPG PROTEIN.
FT   CHAIN      1988   2229       NIA-PRO PROTEIN.
FT   CHAIN      2230   2750       NIB PROTEIN.
FT   CHAIN      2751   3071       COAT PROTEIN.
SQ   SEQUENCE   3071 AA;  349970 MW;  2FD59A0668DE2FB7 CRC64;

  Query Match             13.0%;  Score 70.5;  DB 12;  Length 3071;
  Best Local Similarity   33.3%;  Pred. No. 2.5e+02;
  Matches   22;  Conservative   10;  Mismatches   17;  Indels   17;  Gaps    3;

Qy      22 TGEDNDEFFMDFLQT----LLVGTPEELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMH 77
           ||:  | :|||| |    :|  : | |||| || :| :  |:           ::|:
Db    2363 TGKKKD-YFMDFSDTDKEEILKQSCERLYEGQLGVWNGSLKAE----------IRPIE 2409

Qy      78 KAELVK 83
           | |  |
Db    2410 KTEANK 2415


RESULT 12
Q8QXL1
ID   Q8QXL1      PRELIMINARY;      PRT;  3071 AA.
AC   Q8QXL1;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Polyprotein.
OS   Sorghum mosaic virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC   Potyvirus.
OX   NCBI_TaxID=32619;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YUHANG;
RA   Chen J., Chen J.P., Adams M.J.;
RT   "Characterisation of potyviruses from sugarcane and maize in China.";
RL   Arch. Virol. 147:0-0(2002).
DR   EMBL; AJ310198; CAC84438.1; -.
FT   CHAIN         2    233       P1 PROTEIN.
FT   CHAIN       234    693       HC-PRO PROTEIN.
FT   CHAIN       694   1040       P3 PROTEIN.
FT   CHAIN      1041   1107       6K1 PROTEIN.
FT   CHAIN      1108   1745       CI PROTEIN.
FT   CHAIN      1746   1798       6K2 PROTEIN.
FT   CHAIN      1799   1987       NIA-VPG PROTEIN.
FT   CHAIN      1988   2229       NIA-PRO PROTEIN.
FT   CHAIN      2230   2750       NIB PROTEIN.
FT   CHAIN      2751   3071       COAT PROTEIN.
SQ   SEQUENCE   3071 AA;  349537 MW;  E3D326DE1B9ACB05 CRC64;

  Query Match             13.0%;  Score 70.5;  DB 12;  Length 3071;
  Best Local Similarity   33.3%;  Pred. No. 2.5e+02;
  Matches   22;  Conservative   10;  Mismatches   17;  Indels   17;  Gaps    3;

Qy      22 TGEDNDEFFMDFLQT----LLVGTPEELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMH 77
           ||:  | :|||| |    :|  : | |||| || :| :  |:           ::|:
Db    2363 TGKKKD-YFMDFSDTDKEEILKQSCERLYEGQLGVWNGSLKAE----------IRPIE 2409

Qy      78 KAELVK 83
           | |  |
Db    2410 KTEANK 2415


RESULT 13
Q9FW93
ID   Q9FW93      PRELIMINARY;      PRT;  643 AA.
AC   Q9FW93;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Putative cytochrome P450.
GN   OSJNBA0026L12.20.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
RA   Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA   Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA   Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBa0026L12 genomic sequence.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   EMBL; AC068924; AAG13498.1; -.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Monooxygenase; Oxidoreductase.
SQ   SEQUENCE   643 AA;  70706 MW;  C695CA5DF1A36463 CRC64;

  Query Match             12.8%;  Score 69.5;  DB 10;  Length 643;
  Best Local Similarity   22.4%;  Pred. No. 45;
  Matches   33;  Conservative   19;  Mismatches   44;  Indels   51;  Gaps    6;

Qy       3 GSRALLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTP---EELYE------GTLGKY 53
           |:|||    | : |   :  | |:    : ::   ||  |    :|||      |  |::
Db     296 GNRALTDDELVMLCSEFLVAGTDSTSAALQWIMAELVRNPSIQSKLYEEIKSKTGGGGQH 355

Qy      54 NVNEDAKAAMTELKSCI-DGLQ---PMH-----------------KAELVKLLVQVLG 90
```

```
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003772; AAF56994.1; -.
DR   EMBL; AY061195; AAL28743.1; -.
DR   FlyBase; FBgn0039740; CG7928.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 7.
DR   SMART; SM00355; ZnF_C2H2; 6.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 5.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   457 AA;  53091 MW;  2F5D641877EFD97E CRC64;

  Query Match             13.7%;  Score 74;  DB 5;  Length 457;
  Best Local Similarity   27.2%;  Pred. No. 9.2;
  Matches   28;  Conservative   16;  Mismatches   35;  Indels   24;  Gaps    5;

Qy      16 CICRMATGEDNDEFFMDFLQTLLVGTP-----EELYEGTLG-KYNVNEDAKAAMTELKSC 69
           |||: :     |       : |  || |     | : : | |  | : |::    |  | ||
Db       4 CICQFSVRVPKD------IHTDTVGHPPVLISELVLQCTRGTNYVLTEESS---TICKKC 54

Qy      70 IDGLQPMHKA---------ELVKLLVQVLGSQDGAGTDYKDDD 103
             : |   ||:         |:::|:     |:|   | ||:||
Db      55 CEKLARYHKSIQIARKLRGEILELIHSPYMSKDHKQTSYKEDD 97


RESULT 8
Q9GK67
ID   Q9GK67      PRELIMINARY;      PRT;    90 AA.
AC   Q9GK67;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Lipophilin AL.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LACRIMAL GLAND;
RA   Zhao C., Nguyen T.X., Lehrer R.I.;
RT   "Rabbit Lipophilins.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF308614; AAG42802.1; -.
DR   InterPro; IPR000329; Uteroglobin.
DR   Pfam; PF01099; Uteroglobin; 1.
DR   SMART; SM00096; UTG; 1.
SQ   SEQUENCE   90 AA;  9762 MW;  6D34F9540C1FF742 CRC64;

  Query Match             13.5%;  Score 73;  DB 6;  Length 90;
  Best Local Similarity   27.7%;  Pred. No. 1.5;
  Matches   23;  Conservative   15;  Mismatches   41;  Indels    4;  Gaps    2;

Qy       7 LLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMTEL 66
           |||||| | |    |       | :| :  |    |: :|   | ||:   :| :|  ::
Db       7 LLLVALALGCY--EADAAACPAFVLDSVGFLF--DPKPVYRQKLAKYDAPPEAVEAKLQV 62

Qy      67 KSCIDGLQPMHKAELVKLLVQVL 89
           | | | :    :  :  :| :::

Db      63 KECTDEIDKGKRVLIAAVLTKIV 85


RESULT 9
Q9UT83
ID   Q9UT83      PRELIMINARY;      PRT;    609 AA.
AC   Q9UT83;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative RNA-binding protein.
GN   SPAC343.07.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL109739; CAB52270.1; -.
DR   HSSP; P19339; 2SXL.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   InterPro; IPR003955; RRM_2.
DR   Pfam; PF00076; rrm; 3.
DR   SMART; SM00360; RRM; 2.
DR   SMART; SM00362; RRM_2; 1.
DR   PROSITE; PS50102; RRM; 2.
DR   PROSITE; PS00030; RRM_RNP_1; 1.
SQ   SEQUENCE   609 AA;  68547 MW;  0ADE9A51603AB8F3 CRC64;

  Query Match             13.5%;  Score 73;  DB 3;  Length 609;
  Best Local Similarity   23.5%;  Pred. No. 17;
  Matches   23;  Conservative   20;  Mismatches   27;  Indels   28;  Gaps    5;

Qy      18 CRMATG--EDNDEFFMDFLQTLLVG--------TPEELYE---------GTLGKYNVNE 57
           |:|  |   | | :|  |      ::|        :| |||          |     ::||
Db     248 CQMMPGMLLEQNPQFLYDNPSIFVIGILNLPLKVSPVELYNEFSNHGHILGVAINQSINE 307

Qy      58 D----AKAAMTELKSCIDGLQPMH----KAELVKLLVQ 87
           |    |: |::  :|||: ::  |    :  :::| ::
Db     308 DMTHYAEVAVSTYESCIEIIEKFHAIAYEGSILQLFIK 345


RESULT 10
Q99JZ9
ID   Q99JZ9      PRELIMINARY;      PRT;    500 AA.
AC   Q99JZ9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to signal recognition particle 54 kDa.
GN   SRP54.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005543; AAH05543.1; -.
DR   HSSP; O07347; 2FFH.
DR   MGD; MGI:1346087; Srp54.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR000897; SRP54.
DR   InterPro; IPR004125; SRP54_SPB.
DR   Pfam; PF00448; SRP54; 1.
DR   Pfam; PF02881; SRP54_N; 1.
DR   Pfam; PF02978; SRP_SPB; 1.
DR   ProDom; PD000819; SRP54; 1.
```

```
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK011928; BAB27921.1; -.
DR   EMBL; BC019683; AAH19683.1; -.
DR   HSSP; O07347; 2FFH.
DR   MGD; MGI:1346087; Srp54.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR000897; SRP54.
DR   InterPro; IPR004125; SRP54_SPB.
DR   Pfam; PF00448; SRP54; 1.
DR   Pfam; PF02881; SRP54_N; 1.
DR   Pfam; PF02978; SRP_SPB; 1.
DR   ProDom; PD000819; SRP54; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00300; SRP54; 1.
SQ   SEQUENCE   504 AA;  55720 MW;  79AD58BE6D1E89CA CRC64;

  Query Match           14.1%;  Score 76.5;  DB 11;  Length 504;
  Best Local Similarity  32.4%;  Pred. No. 5.5;
  Matches   36;  Conservative   16;  Mismatches   40;  Indels   19;  Gaps    6;

Qy       2 KGSRALLLVALTLFCICRMATGEDNDEF----FMDFLQTLL-VGTPEELYEGTLGKYNVN 56
           ||  ||  || |   |  : |||  |:|       |:  || :|  | | :       ||
Db     255 KGGGALSAVAATKSPIIFIGTGEHIDDFEPFKTQPFISKLLGMGDIEGLID------KVN 308

Qy      57 E----DAKAAMTELKSCIDGLQPMHK--AELVKL--LVQVLGSQDGAGTDY 99
           |    | :| : :||     |: |::    ::|:    |:||   | |||:
Db     309 ELKLDDNEALIEKLKHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDF 359


RESULT 6
Q8XM24
ID   Q8XM24      PRELIMINARY;      PRT;  2104 AA.
AC   Q8XM24;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Probable alpha-N-acetylglucosaminidase.
GN   CPE0866.
OS   Clostridium perfringens.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1502;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=13 / TYPE A;
RX   PubMed=11792842;
RA   Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA   Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT   "Complete genome sequence of Clostridium perfringens, an anaerobic
RT   flesh-eater.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR   EMBL; AP003188; BAB80572.1; -.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR003961; FN_III.
DR   Pfam; PF00754; F5_F8_type_C; 1.
DR   Pfam; PF00041; fn3; 1.
DR   SMART; SM00060; FN3; 1.
DR   PROSITE; PS00018; EF_HAND; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   2104 AA;  236010 MW;  5FE345659599EED7 CRC64;

  Query Match           14.0%;  Score 76;  DB 16;  Length 2104;
  Best Local Similarity  29.0%;  Pred. No. 38;
  Matches   27;  Conservative   10;  Mismatches   34;  Indels   22;  Gaps    3;

Qy      24 EDNDEF---FMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAE 80
           :|:| |    | | |: ||  : :| ||     ||     |     |          |
Db     721 KDSDAFLYDFADILKQLLANSAQEYYEVMCNAYNNGNGEKFKFVSGKF----------LE 770

Qy      81 LVKLLVQVLGSQ---------DGAGTDYKDDDD 104
           |:||  :|| ::         : | |  || ||
Db     771 LIKLQERVLSTRPEFLIGNWIEDARTMLKDSDD 803


RESULT 7
Q9VAB8
ID   Q9VAB8      PRELIMINARY;      PRT;  457 AA.
AC   Q9VAB8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   CG7928 protein (LD15405P).
GN   CG7928.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
```

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Potential ligand-binding  protein.
GN   RYD5.
OS   Rattus rattus (Black rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10117;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FISCHER; TISSUE=OLFACTORY EPITHELIUM;
RX   MEDLINE=92007724; PubMed=1915264;
RA   Dear T.N., Boehm T., Keverne E.B., Rabbitts T.H.;
RT   "Novel genes for potential ligand-binding proteins in subregions of
RT   the olfactory mucosa.";
RL   EMBO J. 10:2813-2819(1991).
DR   EMBL; X60661; CAA43068.1; -.
DR   InterPro; IPR000329; Uteroglobin.
DR   SMART; SM00096; UTG; 1.
SQ   SEQUENCE   94 AA;  10401 MW;  25A4BBBB4977E247 CRC64;

  Query Match             64.5%;  Score 349.5;  DB 11;  Length 94;
  Best Local Similarity   75.3%;  Pred. No. 3e-31;
  Matches   70;  Conservative   11;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       1 MKGSRALLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAK 60
           |||| | ||||||: ||| :   ||::||||:||||||||||||||||| |||||||: ||
Db       1 MKGSSA-LLVALTVLCICGLTRAEDDNEFFMEFLQTLLVGTPEELYEGPLGKYNVNDMAK 59

Qy      61 AAMTELKSCIDGLQPMHKAELVKLLVQVLGSQD 93
           ||:|||||||| |||:|| :|||||||| :|:
Db      60 AALTELKSCIDELQPVHKEQLVKLLVQVLDAQE 92


RESULT 3
Q8VD96
ID   Q8VD96      PRELIMINARY;      PRT;    96 AA.
AC   Q8VD96;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CC10 protein precursor.
GN   CC10.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Gutierrez-Sagal R., Nieto A.;
RT   "Cloning and sequencing of a cDNA encoding hamster uteroglobin/clara
RT   cell 10 kDa protein.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L37041; AAL31349.1; -.
DR   InterPro; IPR003628; Uteroglbn_sub.
DR   InterPro; IPR000329; Uteroglobin.
DR   Pfam; PF01099; Uteroglobin; 1.
DR   ProDom; PD012475; Uteroglbn_sub; 1.
DR   SMART; SM00096; UTG; 1.
DR   PROSITE; PS00404; UTEROGLOBIN_2; UNKNOWN_1.
KW   Signal.
FT   SIGNAL       1     18       POTENTIAL.
SQ   SEQUENCE   96 AA;  10509 MW;  5EB9CBDD46143389 CRC64;

  Query Match             16.7%;  Score 90.5;  DB 11;  Length 96;
  Best Local Similarity   27.8%;  Pred. No. 0.019;
  Matches   25;  Conservative   17;  Mismatches   43;  Indels    5;  Gaps    2;

Qy       8 LLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMTELK 67
           : :|: :  :|  :   |    |   |: | :|: |  ||  |  ||  | : : |:||
Db       5 ITMAVVMLSVCCSSASSDTCPGFFQVLEFLFMGS-ESSYEAALKFYNPGSDLQDSGTQLK 63

Qy      68 SCIDGLQPMHKAELVKL----LVQVLGSQD 93
           :| |    :  ::||    |   | :||
Db      64 KLVDTLPQKTRMNIMKLSEIILTSPLCNQD 93


RESULT 4
O62571
ID   O62571      PRELIMINARY;      PRT;   478 AA.
AC   O62571;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Stress-responsive protein kinase (PRKSD).
OS   Suberites domuncula (Sponge).
OC   Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC   Hadromerida; Suberitidae; Suberites.
OX   NCBI_TaxID=55567;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mueller W.E.G.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96394688; PubMed=8798342;
RA   Kruse M., Gamulin V., Cetkovic H., Pancer Z., Mueller I.M.,
RA   Mueller W.E.G.;
RT   "Molecular evolution of the Metazoan protein kinase C multigene
RT   family.";
RL   J. Mol. Evol. 43:374-383(1996).
DR   EMBL; Y13101; CAA73555.1; -.
DR   HSSP; P24941; 1HCL.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   ATP-binding; Kinase; Transferase.
SQ   SEQUENCE   478 AA;  54098 MW;  6B38FF5FD88D6332 CRC64;

  Query Match             16.1%;  Score 87.5;  DB 5;  Length 478;
  Best Local Similarity   31.0%;  Pred. No. 0.31;
  Matches   27;  Conservative   18;  Mismatches   25;  Indels   17;  Gaps    4;

Qy      22 TGEDNDEF---FMDFLQTLLVGTPEELYEGT-LGKYNVNEDAKAAMTELKSCIDGLQPMH 77
           | :| |:|   | ||:   ||  |||    | | ::   : ||  :  ||
Db     238 TLKDTDKFSNEFSDFISRCLVKNPEERMSATALLQHKFIKSAK-PVAVLK---------- 286

Qy      78 KAELVKLLVQVLGSQDGAGTDYKDDDD 104
             ||::  :::|  ::|:|:| :||:|
Db     287 --ELIQDAMRILEEEEGSGSDEEDDND 311


RESULT 5
Q9D008
ID   Q9D008      PRELIMINARY;      PRT;   504 AA.
AC   Q9D008;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Signal recognition particle 54 kDa.
GN   SRP54.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 14, 2003, 10:50:14 ; Search time 33 Seconds
(without alignments)
655.604 Million cell updates/sec

Title: US-09-768-826-47
Perfect score: 542
Sequence: 1 MKGSRALLLVALTLFCICRM.....VQVLGSQDGAGTDYKDDDDK 105

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	477	88.0	95	4	Q8TD33	Q8td33 homo sapien
2	349.5	64.5	94	11	Q05702	Q05702 rattus ratt
3	90.5	16.7	96	11	Q8VD96	Q8vd96 mesocricetu
4	87.5	16.1	478	5	O62571	O62571 suberites d
5	76.5	14.1	504	11	Q9D008	Q9d008 mus musculu
6	76	14.0	2104	16	Q8XM24	Q8xm24 clostridium
7	74	13.7	457	5	Q9VAB8	Q9vab8 drosophila
8	73	13.5	90	6	Q9GK67	Q9gk67 oryctolagus
9	73	13.5	609	3	Q9UT83	Q9ut83 schizosacch
10	72.5	13.4	500	11	Q99JZ9	Q99jz9 mus musculu
11	70.5	13.0	3071	12	Q8QXL2	Q8qxl2 sorghum mos
12	70.5	13.0	3071	12	Q8QXL1	Q8qxl1 sorghum mos
13	69.5	12.8	643	10	Q9FW93	Q9fw93 oryza sativ
14	69.5	12.8	2458	2	O51827	O51827 pseudomonas
15	69	12.7	540	16	Q9CFY4	Q9cfy4 lactococcus
16	68.5	12.6	2100	3	P87112	P87112 schizosacch
17	68	12.5	1425	4	O95247	O95247 homo sapien
18	68	12.5	1483	4	O95277	O95277 homo sapien
19	68	12.5	1527	4	Q9UIG0	Q9uig0 homo sapien
20	68	12.5	2258	12	Q81081	Q81081 human parai
21	67.5	12.5	249	17	Q8TY89	Q8ty89 methanopyru
22	67.5	12.5	299	5	Q19742	Q19742 caenorhabdi
23	67.5	12.5	317	10	Q8VXX7	Q8vxx7 arabidopsis
24	67.5	12.5	322	16	Q9HVL5	Q9hvl5 pseudomonas
25	67.5	12.5	454	11	Q8QZR1	Q8qzr1 mus musculu
26	67.5	12.5	485	10	Q9LVG2	Q9lvg2 arabidopsis
27	67.5	12.5	1953	3	O13450	O13450 saccharomyc
28	67	12.4	511	10	O80740	O80740 arabidopsis
29	66.5	12.3	358	16	Q8YCG5	Q8ycg5 brucella me
30	66	12.2	151	10	Q8RYJ9	Q8ryj9 oryza sativ
31	66	12.2	189	2	O85808	O85808 borrelia bu
32	66	12.2	192	2	P94238	P94238 borrelia bu
33	66	12.2	210	2	Q44978	Q44978 borrelia bu
34	66	12.2	742	6	Q95JY6	Q95jy6 macaca fasc
35	65.5	12.1	90	6	Q9GK65	Q9gk65 oryctolagus
36	65.5	12.1	362	17	O57942	O57942 pyrococcus
37	65.5	12.1	454	6	Q9XSW4	Q9xsw4 mustela vis
38	65.5	12.1	795	5	O62006	O62006 branchiosto
39	65.5	12.1	1321	16	O86653	O86653 streptomyce
40	65.5	12.1	1687	11	O35651	O35651 mus musculu
41	65.5	12.1	1706	11	P97790	P97790 mus musculu
42	65.5	12.1	1719	11	P97789	P97789 mus musculu
43	65	12.0	91	6	Q9GK66	Q9gk66 oryctolagus
44	65	12.0	158	17	Q97C56	Q97c56 thermoplasm
45	65	12.0	262	5	Q9BPS1	Q9bps1 bombyx mori

ALIGNMENTS

RESULT 1
Q8TD33
ID Q8TD33 PRELIMINARY; PRT; 95 AA.
AC Q8TD33;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative ligand binding protein RYD5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bingle C.D.;
RT "Human RYD5, a new secretoglobin.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY026938; AAK08972.1; -.
SQ SEQUENCE 95 AA; 10457 MW; F629AF06C96D2392 CRC64;

Query Match 88.0%; Score 477; DB 4; Length 95;
Best Local Similarity 98.9%; Pred. No. 2.1e-45;
Matches 94; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 MKGSRALLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAK 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKGSRALLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAK 60

Qy      61 AAMTELKSCIDGLQPMHKAELVKLLVQVLGSQDGA 95
           ||||||||| ||||||||||||||||||||||||||
Db      61 AAMTELKSCRDGLQPMHKAELVKLLVQVLGSQDGA 95
```

RESULT 2
Q05702
ID Q05702 PRELIMINARY; PRT; 94 AA.
AC Q05702;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```
  Best Local Similarity   32.1%;  Pred. No. 1.2e+02;
  Matches   17;  Conservative   14;  Mismatches   19;  Indels    3;  Gaps    2;

Qy      44 ELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGSQDGAG 96
           :|   |  :  : :|| || | ::    ||||: ::::|: : :||| |'  |
Db    2239 QLLHLTQSEQELEQDA-AARTAMQG--SGLQPLQRSQIVQAIARVLGGQGQCG 2288


RESULT 15
G86790
fibronectin-binding protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: G86790
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: G86790
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-540 <STO>
A;Cross-references: GB:AE005176; PID:g12724308; PIDN:AAK05425.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yngB

  Query Match             12.7%;  Score 69;  DB 2;  Length 540;
  Best Local Similarity   30.6%;  Pred. No. 25;
  Matches   22;  Conservative    9;  Mismatches   37;  Indels    4;  Gaps    2;

Qy      25 DNDEFFM---DFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMTELKSCIDGLQP-MHKAE 80
           || | |    : | | |   | :    ||  |  ||  : |:    | :  |   |: :
Db     294 DNAEIFRQKGELLNTFLNQVPNDKTSVTLENYYTNEPIEIALNPALSPVQNAQRYFHRYQ 353

Qy      81 LVKLLVQVLGSQ 92
           :|  |: || |
Db     354 KLKQAVKFLGEQ 365


Search completed: May 14, 2003, 10:53:02
Job time : 19 secs
```

```
R;Bernstein, H.
submitted to the EMBL Data Library, November 1989
A;Reference number: S14553
A;Accession: S14553
A;Molecule type: mRNA
A;Residues: 1-42,'DV',45-504 <BER2>
A;Cross-references: EMBL:X16319; NID:g54193; PIDN:CAA34386.1; PID:g54194
C;Superfamily: signal recognition particle 54K protein
F;2-504/Product: signal recognition particle 54K protein #status predicted <MAT>

  Query Match            14.1%;  Score 76.5;  DB 2;  Length 504;
  Best Local Similarity  32.4%;  Pred. No. 3.7;
  Matches   36;  Conservative   16;  Mismatches   40;  Indels   19;  Gaps    6;

Qy       2 KGSRALLLVALTLFCICRMATGEDNDEF----FMDFLQTLL-VGTPEELYEGTLGKYNVN 56
           ||  ||  || |    |  : |||  |:|       |:  || :|  | | :          ||
Db     255 KGGGALSAVAATKSPIIFIGTGEHIDDFEPFKTQPFISKLLGMGDIEGLID------KVN 308

Qy      57 E----DAKAAMTELKSCIDGLQPMHK--AELVKL--LVQVLGSQDGAGTDY 99
           |    | :| : :||     |: |::   ::|:   |:||   | |||:
Db     309 ELKLDDNEALIEKLKHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDF 359


RESULT 11
S54143
SRP 54 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 23-Jul-1999
C;Accession: S54143
R;Patel, S.
submitted to the EMBL Data Library, April 1995
A;Reference number: S54143
A;Accession: S54143
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-504 <PAT>
A;Cross-references: EMBL:X86373; NID:g784998; PIDN:CAA60132.1; PID:g784999
C;Superfamily: signal recognition particle 54K protein

  Query Match            14.1%;  Score 76.5;  DB 2;  Length 504;
  Best Local Similarity  32.4%;  Pred. No. 3.7;
  Matches   36;  Conservative   16;  Mismatches   40;  Indels   19;  Gaps    6;

Qy       2 KGSRALLLVALTLFCICRMATGEDNDEF----FMDFLQTLL-VGTPEELYEGTLGKYNVN 56
           ||  ||  || |    |  : |||  |:|       |:  || :|  | | :          ||
Db     255 KGGGALSAVAATKSPIIFIGTGEHIDDFEPFKTQPFISKLLGMGDIEGLID------KVN 308

Qy      57 E----DAKAAMTELKSCIDGLQPMHK--AELVKL--LVQVLGSQDGAGTDY 99
           |    | :| : :||     |: |::   ::|:   |:||   | |||:
Db     309 ELKLDDNEALIEKLKHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDF 359


RESULT 12
T38656
probable RNA-binding protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T38656
R;Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z21804
A;Accession: T38656
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-609 <MUR>
A;Cross-references: EMBL:AL109739; NID:e1534774; PIDN:CAB52270.1; GSPDB:GN00066; SPDB:SP
A;Experimental source: strain 972h-; cosmid c343
C;Genetics:
A;Gene: SPDB:SPAC343.07
A;Map position: 1
A;Introns: 44/1; 96/2; 325/3

  Query Match            13.5%;  Score 73;  DB 2;  Length 609;
  Best Local Similarity  23.5%;  Pred. No. 11;
  Matches   23;  Conservative   20;  Mismatches   27;  Indels   28;  Gaps    5;

Qy      18 CRMATG---EDNDEFFMDFLQTLLVG--------TPEELYE---------GTLGKYNVNE 57
           |:|  |    | | :|  |     ::|         :| |||           |      ::||
Db     248 CQMMPGMLLEQNPQFLYDNPSIFVIGILNLPLKVSPVELYNEFSNHGHILGVAINQSINE 307

Qy      58 D----AKAAMTELKSCIDGLQPMH----KAELVKLLVQ 87
           |    |: |::  :|||: ::  |    :  :::| ::
Db     308 DMTHYAEVAVSTYESCIEIIEKFHAIAYEGSILQLFIK 345


RESULT 13
S41122
mannose-6-phosphate isomerase (EC 5.3.1.8) - human
N;Alternate names: phosphomannose isomerase
C;Species: Homo sapiens (man)
C;Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999
C;Accession: S41122; S38666
R;Proudfoot, A.E.I.; Turcatti, G.; Wells, T.N.C.; Payton, M.A.; Smith, D.J.
Eur. J. Biochem. 219, 415-423, 1994
A;Title: Purification, cDNA cloning and heterologous expression of human phosphomanno
A;Reference number: S41122; MUID:94139717; PMID:8307007
A;Accession: S41122
A;Molecule type: mRNA
A;Residues: 1-423 <PRO>
A;Cross-references: EMBL:X76057; NID:g416016; PIDN:CAA53657.1; PID:g416017
C;Genetics:
A;Gene: GDB:MPI
A;Cross-references: GDB:119397; OMIM:154550
A;Map position: 15q22-15qter
C;Superfamily: yeast mannose-6-phosphate isomerase
C;Keywords: intramolecular oxidoreductase; isomerase

  Query Match            13.0%;  Score 70.5;  DB 2;  Length 423;
  Best Local Similarity  25.7%;  Pred. No. 13;
  Matches   27;  Conservative   19;  Mismatches   38;  Indels   21;  Gaps    5;

Qy       8 LLVALTLF-CICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLG-------KYNVNEDA 59
           : :||| |  :|      |:        : | |   ||  ::  :|        |  :: |:
Db     138 MAIALTPFQGLCGFRPVEE--------IVTFLKKVPE--FQFLIGDEAATHLKQTMSHDS 187

Qy      60 KAAMTELKSCIDGLQPMHK---AELVKLLVQVLGSQDGAGTDYKD 101
           :|  : |:||   |    |    | : |||: :  |  || : :|
Db     188 QAVASSLQSCFSHLMKSEKKVVVEQLNLLVKRISQQAAAGNNMED 232


RESULT 14
T17420
probable polyketide synthase type I - Pseudomonas fluorescens
C;Species: Pseudomonas fluorescens
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000
C;Accession: T17420
R;Nowak-Thompson, B.; Chaney, N.; Wing, J.S.; Gould, S.J.; Loper, J.E.
J. Bacteriol. 181, 2166-2174, 1999
A;Title: Characterization of the pyoluteorin biosynthetic gene cluster of Pseudomonas
A;Reference number: Z18776; MUID:99194726; PMID:10094695
A;Accession: T17420
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2458 <NOW>
A;Cross-references: EMBL:AF081920; NID:g4582974; PID:g2781416; PIDN:AAC38075.1
C;Genetics:
A;Gene: pltB
C;Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier pro
C;Keywords: carrier protein
F;31-429/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F;535-815/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F;939-1009/Domain: acyl carrier protein homology <ACP1>
F;1053-1446/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F;2337-2408/Domain: acyl carrier protein homology <ACP2>

  Query Match            12.8%;  Score 69.5;  DB 2;  Length 2458;
```

```
A;Cross-references: GB:L04503; NID:g202313; PIDN:AAA03625.1; PID:g433093
C;Genetics:
A;Introns: 19/1; 81/3
C;Complex: homodimer linked by two disulfide bonds
C;Superfamily: uteroglobin
C;Keywords: lung; steroid binding; uterus
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-96/Product: uteroglobin #status predicted <MAT>
F;24/Disulfide bonds: interchain (to 90) #status predicted
F;90/Disulfide bonds: interchain (to 24) #status predicted

  Query Match             15.0%;  Score 81.5;  DB 1;  Length 96;
  Best Local Similarity   26.2%;  Pred. No. 0.17;
  Matches   22;  Conservative   19;  Mismatches   42;  Indels    1;  Gaps    1;

Qy       8 LLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMTELK 67
           : : : :  ||  :   |    |:  |: ||: : |  |  :|: :|  | |  | |:||
Db       5 ITITVVMLSICCSSASSDICPGFLQVLEALLMES-ESGYVASLKPFNPGSDLQNAGTQLK 63

Qy      68 SCIDGLQPMHKAELVKLLVQVLGS 91
             :| |    :  ::||  ::| |
Db      64 RLVDTLPQETRINIMKLTEKILTS 87


RESULT 7
UGRBL
uteroglobin precursor - brown hare
N;Alternate names: blastokinin
C;Species: Lepus capensis (brown hare)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 22-Jun-1999
C;Accession: A23825
R;Lopez de Haro, M.S.; Nieto, A.
Biochem. J. 235, 895-898, 1986
A;Title: Nucleotide and derived amino acid sequences of a cDNA coding for pre-uteroglobi
A;Reference number: A23825; MUID:86323069; PMID:3019311
A;Accession: A23825
A;Molecule type: mRNA
A;Residues: 1-91 <LOP>
A;Cross-references: GB:M25609; NID:g164246; PIDN:AAA30960.1; PID:g164247
A;Experimental source: lung
C;Comment: Uteroglobin, synthesized in the uterus and lung, is secreted by the uterus up
C;Complex: homodimer linked by two disulfide bonds
C;Superfamily: uteroglobin
C;Keywords: lung; steroid binding; uterus
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-91/Product: uteroglobin #status predicted <MAT>
F;24/Disulfide bonds: interchain (to 90) #status predicted
F;90/Disulfide bonds: interchain (to 24) #status predicted

  Query Match             14.8%;  Score 80;  DB 1;  Length 91;
  Best Local Similarity   26.8%;  Pred. No. 0.23;
  Matches   26;  Conservative   18;  Mismatches   37;  Indels   16;  Gaps    3;

Qy       1 MKGSRALLLVALTLFC------ICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYN 54
           || :  | || | | |      ||           |   :: ||:|||   | :| ::
Db       1 MKLTITLALVTLALLCSPASAGIC---------PGFAHVIENLLLGTPSS-YGTSLKEFQ 50

Qy      55 VNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGS 91
            ::  | |  ::|  :| |    :  ::||. ::: |
Db      51 PDDAMKDAGMQMKKVLDTLPQTTRENIIKLTEKIVKS 87


RESULT 8
JC2026
cell specific 10K protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 17-Mar-1999
C;Accession: JC2026
R;Ray, M.K.; Magdaleno, S.; O'Malley, B.W.; DeMayo, F.J.
Biochem. Biophys. Res. Commun. 197, 163-171, 1993
A;Title: Cloning and characterization of the mouse clara cell specific 10 kDa protein ge
A;Reference number: JC2026; MUID:94071937; PMID:7916613
A;Accession: JC2026
A;Molecule type: DNA
A;Residues: 1-113 <RAY>
C;Comment: This protein is the major secretory product of the Clara cell and binds to
C;Superfamily: uteroglobin
F;73/Region: ochre stop codon

  Query Match             14.5%;  Score 78.5;  DB 2;  Length 113;
  Best Local Similarity   30.7%;  Pred. No. 0.43;
  Matches   27;  Conservative   15;  Mismatches   43;  Indels    3;  Gaps    2;

Qy       6 ALLLVALTLFCICRMATGEDND--EFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAAM 63
             :||   |  ||    ||   :|    |:  |: ||: : |  |  :|  :|   |   |
Db      18 SLLEKKLLSFCFFLPATPASSDICPGFLQVLEALLMES-ESGYVASLKPFNPGSDLXNAG 76

Qy      64 TELKSCIDGLQPMHKAELVKLLVQVLGS 91
           |:||  :| |    :  ::||  ::| |
Db      77 TQLKRLVDTLPQETRINIMKLTEKILTS 104


RESULT 9
S05197
signal recognition particle 54K protein - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 02-Feb-2001
C;Accession: S05197; S10619
R;Roemisch, K.; Webb, J.; Herz, J.; Prehn, S.; Frank, R.; Vingron, M.; Dobberstein, B
Nature 340, 478-482, 1989
A;Title: Homology of 54K protein of signal-recognition particle, docking protein and
A;Reference number: S05197; MUID:89330596; PMID:2502717
A;Accession: S05197
A;Molecule type: mRNA
A;Residues: 1-504 <ROE>
A;Cross-references: EMBL:X16318; NID:g928; PIDN:CAA34385.1; PID:g930
A;Note: part of this sequence, including the amino end of the mature protein, was con
R;Bernstein, H.D.; Poritz, M.A.; Strub, K.; Hoben, P.J.; Brenner, S.; Walter, P.
Nature 340, 482-486, 1989
A;Title: Model for signal sequence recognition from amino-acid sequence of 54K subuni
A;Reference number: S05198; MUID:89330597; PMID:2502718
A;Accession: S10619
A;Molecule type: protein
A;Residues: 2-35;58-70;120-126;141-153;374-388;416-426;470-485 <BER>
C;Superfamily: signal recognition particle 54K protein
C;Keywords: GTP binding; nucleotide binding; P-loop
F;1-504/Product: signal recognition particle 54K protein #status experimental <MAT>
F;108-115/Region: nucleotide-binding motif A (P-loop)

  Query Match             14.1%;  Score 76.5;  DB 2;  Length 504;
  Best Local Similarity   32.4%;  Pred. No. 3.7;
  Matches   36;  Conservative   16;  Mismatches   40;  Indels   19;  Gaps    6;

Qy       2 KGSRALLLVALTLFCICRMATGEDNDEF----FMDFLQTLL-VGTPEELYEGTLGKYNVN 56
           ||  ||  || |   |  : |||  |:|       |:  || :|  | | :       ||
Db     255 KGGGALSAVAATKSPIIFIGTGEHIDDFEPFKTQPFISKLLGMGDIEGLID------KVN 308

Qy      57 E----DAKAAMTELKSCIDGLQPMHK--AELVKL--LVQVLGSQDGAGTDY 99
           |    | :| : :||     |: |::    ::|:    |:||   | |||:
Db     309 ELKLDDNEALIEKLKHGQFTLRDMYEQFQNIMKMGPFSQILGMIPGFGTDF 359


RESULT 10
S05198
signal recognition particle 54K protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 23-Jul-1999
C;Accession: S05198; S14553
R;Bernstein, H.D.; Poritz, M.A.; Strub, K.; Hoben, P.J.; Brenner, S.; Walter, P.
Nature 340, 482-486, 1989
A;Title: Model for signal sequence recognition from amino-acid sequence of 54K subuni
A;Reference number: S05198; MUID:89330597; PMID:2502718
A;Accession: S05198
A;Molecule type: mRNA
A;Residues: 1-504 <BER1>
A;Cross-references: EMBL:X16319
```

A;Reference number: A24217; MUID:86056319; PMID:2415398
A;Accession: A24217
A;Molecule type: mRNA
A;Residues: 22-91 <LOP>
A;Cross-references: GB:M27564; NID:g165792; PIDN:AAA31496.1; PID:g165793
A;Experimental source: lung
R;Atger, M.; Mercier, J.C.; Haze, G.; Fridlansky, F.; Milgrom, E.
Biochem. J. 177, 985-988, 1979
A;Title: N-terminal sequences of uteroglobin and its precursor.
A;Reference number: A90303; MUID:79187160; PMID:571719
A;Accession: A90303
A;Molecule type: protein
A;Residues: 1-5,'F',7-10,'X',15,'G',17-54,'X',56,'B',58-66,'B',68-70,'XX',73 <ATG>
R;Ponstingl, H.; Nieto, A.; Beato, M.
Biochemistry 17, 3908-3912, 1978
A;Title: Amino acid sequence of progesterone-induced rabbit uteroglobin.
A;Reference number: A90417; MUID:79042086; PMID:568483
A;Accession: A90417
A;Molecule type: protein
A;Residues: 22-81,'Q',83-91 <PON>
R;Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.
Proc. Natl. Acad. Sci. U.S.A. 75, 5516-5519, 1978
A;Title: Amino acid sequence of a progesterone-binding protein!
A;Reference number: A93824; MUID:79074850; PMID:281700
A;Accession: A93824
A;Molecule type: protein
A;Residues: 22-49,'D',51,'EN',54-59,61-66,'NEPSL',72-91 <POP>
R;Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.
submitted to the Atlas, October 1982
A;Reference number: A94608
A;Accession: A94608
A;Molecule type: protein
A;Residues: 50-62;67-71 <PO2>
R;Morize, I.; Surcouf, E.; Vaney, M.C.; Buehner, M.; Mornon, J.P.
submitted to the Brookhaven Protein Data Bank, April 1989
A;Reference number: A50025; PDB:1UTG
A;Contents: annotation; X-ray crystallography, 1.34 angstroms, residues 22-91
R;Morize, I.; Surcouf, E.; Vaney, M.C.; Epelboin, Y.; Buehner, M.; Fridlansky, F.; Milgr
J. Mol. Biol. 194, 725-739, 1987
A;Title: Refinement of the C222-1 crystal form of oxidized uteroglobin at 1.34 angstroms
A;Reference number: A44652; MUID:88011213; PMID:3656405
A;Contents: annotation; X-ray crystallography, 1.34 angstroms
R;Bally, R.; Delettre, J.
submitted to the Brookhaven Protein Data Bank, May 1989
A;Reference number: A50553; PDB:2UTG
A;Contents: annotation; X-ray crystallography, 1.64 angstroms, residues 22-91
R;Bally, R.; Delettre, J.
J. Mol. Biol. 206, 153-170, 1989
A;Title: Structure and refinement of the oxidized P2-1 form of uteroglobin at 1.64 angst
A;Reference number: A44653; MUID:89199637; PMID:2704039
A;Contents: annotation; X-ray crystallography, 1.64 angstroms; disulfide bonds
R;Menne, C.; Suske, G.; Arnemann, J.; Wenz, M.; Cato, A.C.B.; Beato, M.
Proc. Natl. Acad. Sci. U.S.A. 79, 4853-4857, 1982
A;Title: Isolation and structure of the gene for the progesterone-inducible protein uter
A;Reference number: I46904; MUID:83014990; PMID:6956897
A;Accession: I46904
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-91 <MEN>
A;Cross-references: GB:J00689; NID:g165786; PIDN:AAA31495.1; PID:g165788
R;Chandra, T.; Woo, S.L.C.; Bullock, D.W.
Biochem. Biophys. Res. Commun. 95, 197-204, 1980
A;Title: Cloning of the rabbit uteroglobin structural gene.
A;Reference number: I46905; MUID:81021016; PMID:7417250
A;Accession: I46905
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 39-77 <CH2>
A;Cross-references: GB:M25057; NID:g165802; PIDN:AAA31498.1; PID:g165803
R;Suske, G.; Menne, C.; Cato, A.; Wenz, M.; Beato, M.
Prog. Clin. Biol. Res. 85, 139-146, 1982
A;Title: Structure and regulated expression of the uteroglobin gene.
A;Reference number: I46907; MUID:82275176; PMID:6287481
A;Accession: I46907
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-45,'V',47-91 <SU2>
A;Cross-references: GB:M32012; NID:g165807; PIDN:AAA31500.1; PID:g165809
R;Atger, M.; Perricaudet, M.; Tiollais, P.; Milgrom, E.
Biochem. Biophys. Res. Commun. 93, 1082-1088, 1980
A;Title: Bacterial cloning of the rabbit uteroglobin structural gene.
A;Reference number: I46906; MUID:80241888; PMID:6156676
A;Accession: I46906
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 53-66,'NT',69-72 <AT2>
A;Cross-references: GB:M25038; NID:g165804; PIDN:AAA31499.1; PID:g165805
C;Comment: Uteroglobin is secreted by the uterus upon induction by progesterone. It b
C;Genetics:
A;Introns: 19/1; 81/3
C;Complex: homodimer linked by two disulfide bonds
C;Superfamily: uteroglobin
C;Keywords: homodimer; steroid binding; uterus
F;1-21/Domain: signal sequence #status experimental <SIG>
F;22-91/Product: uteroglobin #status experimental <MAT>
F;24/Disulfide bonds: interchain (to 90) #status experimental
F;90/Disulfide bonds: interchain (to 24) #status experimental

Query Match 16.4%; Score 89; DB 1; Length 91;
Best Local Similarity 27.8%; Pred. No. 0.026;
Matches 27; Conservative 18; Mismatches 36; Indels 16; Gaps 3;

```
Qy       1 MKGSRALLLVALTLFC------ICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYN 54
           || :  | || | | |      ||           |   :: ||:|||   || :| ::
Db       1 MKLAITLALVTLALLCSPASAGICPR---------FAHVIENLLLGTPSS-YETSLKEFE 50

Qy      55 VNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGS 91
             ::  | |  ::|  :| |    :  ::||  ::: |
Db      51 PDDTMKDAGMQMKKVLDSLPQTTRENIMKLTEKIVKS 87
```

RESULT 6
UGMS
uteroglobin precursor - mouse
N;Alternate names: CC10; Clara cell 10K protein precursor; Clara cell secretory prote
C;Species: Mus musculus (house mouse)
C;Date: 03-May-1994 #sequence_revision 21-Jan-1997 #text_change 22-Jun-1999
C;Accession: A53025; A56656; I51925; S24783
R;Stripp, B.R.; Huffman, J.A.; Bohinski, R.J.
Genomics 20, 27-35, 1994
A;Title: Structure and regulation of the murine Clara cell secretory protein gene.
A;Reference number: A53025; MUID:94292183; PMID:8020953
A;Accession: A53025
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-96 <STR>
A;Cross-references: GB:L24372; NID:g461147; PIDN:AAA65446.1; PID:g785054
R;Singh, G.; Katyal, S.L.; Brown, W.E.; Kennedy, A.L.
Exp. Lung Res. 19, 67-75, 1993
A;Title: Mouse Clara cell 10-kDa (CC10) protein: cDNA nucleotide sequence and molecul
A;Reference number: A56656; MUID:93178380; PMID:8440203
A;Accession: A56656
A;Molecule type: mRNA; protein
A;Residues: 1-96 <SIN>
A;Cross-references: EMBL:X67702; NID:g49690; PIDN:CAA47936.1; PID:g49691
A;Experimental source: lung
A;Note: sequence extracted from NCBI backbone (NCBIP:126148)
A;Note: parts of this sequence, including the amino end of the mature protein, were c
R;Margraf, L.R.; Finegold, M.J.; Stanley, L.A.; Major, A.; Nawkins, H.K.; DeMayo, F.J
Am. J. Respir. Cell Mol. Biol. 9, 231-238, 1993
A;Title: Cloning and tissue-specific expression of the cDNA for the mouse Clara cell
A;Reference number: I51925; MUID:94000840; PMID:8398159
A;Accession: I51925
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-96 <RES>

```
A;Residues: 1-18 <HAG>
A;Cross-references: EMBL:X51318; NID:g55536; PIDN:CAA35701.1; PID:g55537
R;Umland, T.C.; Swaminathan, S.; Furey, W.; Singh, G.; Pletcher, J.; Sax, M.
J. Mol. Biol. 224, 441-448, 1992
A;Title: Refined structure of rat Clara cell 17 kDa protein at 3.0 A resolution.
A;Reference number: S21676; MUID:92219263; PMID:1560460
A;Contents: annotation, X-ray crystallography, 3.0 angstroms
C;Superfamily: uteroglobin
F;24/Disulfide bonds: interchain (to 90) #status experimental
F;90/Disulfide bonds: interchain (to 24) #status experimental

  Query Match             17.8%;  Score 96.5;  DB 2;  Length 96;
  Best Local Similarity   29.8%;  Pred. No. 0.0044;
  Matches   25;  Conservative   17;  Mismatches   41;  Indels    1;  Gaps    1;

Qy       8 LLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMTELK 67
           : : : :  ||  :   |    |:  |: ||:|: |  ||  |  :| . | : | |:||
Db       5 ITITVLMLSICCSSASSDICPGFLQVLEALLLGS-ESNYEAALKPFNPASDLQNAGTQLK 63

Qy      68 SCIDGLQPMHKAELVKLLVQVLGS 91
             :| |    :  :|||  ::| |
Db      64 RLVDTLPQETRINIVKLTEKILTS 87


RESULT 3
JS0036
Clara cell 10K protein precursor - human
N;Alternate names: urinary protein 1
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Aug-1999
C;Accession: JS0036; PS0309; A56890; I38397
R;Singh, G.; Katyal, S.L.; Brown, W.E.; Philips, S.; Kennedy, A.L.; Anthony, J.; Squegli
Biochim. Biophys. Acta 950, 329-337, 1988
A;Title: Amino-acid and cDNA nucleotide sequences of human Clara cell 10kDa protein.
A;Reference number: JS0036; MUID:89000784; PMID:3167058
A;Accession: JS0036
A;Molecule type: mRNA
A;Residues: 1-91 <SIN>
A;Cross-references: GB:X13197; NID:g23131; PIDN:CAA31584.1; PID:g23132
A;Accession: PS0309
A;Molecule type: protein
A;Residues: 22-23,'X',25-28,'X',30-31,'X',33-36 <SI2>
R;Bernard, A.; Roels, H.; Lauwerys, R.; Witters, R.; Gielens, C.; Soumillion, A.; Van Da
Clin. Chim. Acta 207, 239-249, 1992
A;Title: Human urinary protein 1: evidence for identity with the Clara cell protein and
A;Reference number: A56890; MUID:93009001; PMID:1395029
A;Accession: A56890
A;Molecule type: protein
A;Residues: 22-45 <BER>
A;Experimental source: urine
A;Note: sequence extracted from NCBI backbone (NCBIP:119391)
R;Hay, J.G.; Danel, C.; Chu, C.; Crystal, R.G.
Am. J. Physiol. 268, 565-575, 1995
A;Title: Human CC10 gene expression in airway epithelium and subchromosomal locus sugges
A;Reference number: I38397
A;Accession: I38397
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-91 <RES>
A;Cross-references: EMBL:U01101; NID:g457932; PIDN:AAA81885.1; PID:g457933
C;Comment: This protein consists of two identical polypeptides linked by two disulfide b
C;Genetics:
A;Gene: CC10
C;Superfamily: uteroglobin
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-91/Product: Clara cell 10K protein #status experimental <MAT>

  Query Match             17.0%;  Score 92;  DB 2;  Length 91;
  Best Local Similarity   27.8%;  Pred. No. 0.012;
  Matches   27;  Conservative   18;  Mismatches   36;  Indels   16;  Gaps    3;

Qy       1 MKGSRALLLVALTLFC------ICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYN 54
           ||  :  | || | | |      ||           |   ::|||: ||  ||  :  ::

Db       1 MKLAVTLTLVTLALCCSSASAEIC---------PSFQRVIETLLMDTPSS-YEAAMELFS 50

Qy      55 VNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGS 91
             ::| : |  :||  :| |    :  ::||: ::  |
Db      51 PDQDMREAGAQLKKLVDTLPQKPRESIIKLMEKIAQS 87


RESULT 4
S68231
FHG22 protein precursor - golden hamster
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 21-Aug-1998
C;Accession: S68231
R;Dominguez, P.
FEBS Lett. 376, 257-261, 1995
A;Title: Cloning of a Syrian hamster cDNA related to sexual dimorphism: establishment
A;Reference number: S68231; MUID:96105393; PMID:7498554
A;Accession: S68231
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-95 <DOM>
A;Cross-references: EMBL:Z66540
C;Superfamily: uteroglobin
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-95/Product: FHG22 protein #status predicted <MAT>

  Query Match             16.8%;  Score 91;  DB 2;  Length 95;
  Best Local Similarity   31.6%;  Pred. No. 0.017;
  Matches   31;  Conservative   13;  Mismatches   36;  Indels   18;  Gaps    4;

Qy       1 MKGSRALLLVALTLFC-------ICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKY 53
           || |  |||| | : |         :|  |     :  |  | :              | :|  |  :
Db       1 MKLSLCLLLVILAVHCYEANAANVC-PAVLSVSKSFLFDKV---------EKFEAYLQTF 50

Qy      54 NVNEDAKAAMTELKSCIDG-LQPMHKAELVKLLVQVLG 90
             |   :|  |  |:| |||  |  : | |: |:| :|:|
Db      51 NAPPEAVKAKVEVKKCIDSTLNYLEKMEMGKILAEVVG 88


RESULT 5
UGRB
uteroglobin precursor [validated] - rabbit
N;Alternate names: blastokinin
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 28-Feb-1980 #sequence_revision 15-Oct-1982 #text_change 15-Sep-2000
C;Accession: A92391; A93461; A90935; A24217; A90303; A90417; A93824; A94608; I46904;
R;Bailly, A.; Atger, M.; Atger, P.; Cerbon, M.A.; Alizon, M.; Vu Hai, M.T.; Logeat, F
J. Biol. Chem. 258, 10384-10389, 1983
A;Title: The rabbit uteroglobin gene. Structure and interaction with the progesterone
A;Reference number: A92391; MUID:83290960; PMID:6309802
A;Accession: A92391
A;Molecule type: DNA
A;Residues: 1-91 <BAI>
A;Cross-references: GB:K00049; NID:g165789
R;Suske, G.; Wenz, M.; Cato, A.C.B.; Beato, M.
Nucleic Acids Res. 11, 2257-2271, 1983
A;Title: The uteroglobin gene region: hormonal regulation, repetitive elements and co
A;Reference number: A93461; MUID:83220783; PMID:6304644
A;Accession: A93461
A;Molecule type: DNA
A;Residues: 1-91 <SUS>
A;Cross-references: GB:J00687; NID:g1772; PIDN:CAA25669.1; PID:g313668
R;Chandra, T.; Bullock, D.W.; Woo, S.L.C.
DNA 1, 19-26, 1981
A;Title: Hormonally regulated mammalian gene expression: steady-state level and nucle
A;Reference number: A90935; MUID:83157105; PMID:6299663
A;Accession: A90935
A;Molecule type: mRNA
A;Residues: 1-91 <CHA>
A;Cross-references: GB:K01657; NID:g165794; PIDN:AAA31497.1; PID:g165795
R;Lopez de Haro, M.S.; Nieto, A.
FEBS Lett. 193, 247-249, 1985
A;Title: Primary structure of rabbit lung uteroglobin as deduced from the nucleotide
```

AC NO: S17449 ; Database: PIR_73

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 14, 2003, 10:50:34 ; Search time 17 Seconds
(without alignments)
593.771 Million cell updates/sec

Title: US-09-768-826-47
Perfect score: 542
Sequence: 1 MKGSRALLLVALTLFCICRM.....VQVLGSQDGAGTDYKDDDDK 105

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	349.5	64.5	94	2	S17449	probable ligand-bi
2	96.5	17.8	96	2	A36581	polychlorinated bi
3	92	17.0	91	2	JS0036	Clara cell 10K pro
4	91	16.8	95	2	S68231	FHG22 protein prec
5	89	16.4	91	1	UGRB	uteroglobin precur
6	81.5	15.0	96	1	UGMS	uteroglobin precur
7	80	14.8	91	1	UGRBL	uteroglobin precur
8	78.5	14.5	113	2	JC2026	cell specific 10K
9	76.5	14.1	504	2	S05197	signal recognition
10	76.5	14.1	504	2	S05198	signal recognition
11	76.5	14.1	504	2	S54143	SRP 54 protein - h
12	73	13.5	609	2	T38656	probable RNA-bindi
13	70.5	13.0	423	2	S41122	mannose-6-phosphat
14	69.5	12.8	2458	2	T17420	probable polyketid
15	69	12.7	540	2	G86790	fibronectin-bindin
16	68.5	12.6	2100	2	T38128	t7i23.15 protein -
17	68	12.5	92	2	A56413	major allergen Fel
18	68	12.5	270	2	S30947	hypothetical prote
19	67.5	12.5	299	2	T16137	hypothetical prote
20	67.5	12.5	322	2	C83075	octaprenyl-diphosp
21	67.5	12.5	1953	2	S63244	BNI1 protein - yea
22	67	12.4	511	2	T02269	hypothetical prote
23	66.5	12.3	358	2	AD3580	iron(III)-binding
24	66	12.2	192	2	S70285	outer surface prot
25	65.5	12.1	92	2	JC1136	major allergen cha
26	65.5	12.1	122	2	A43644	sarcocystatin A pr
27	65.5	12.1	362	2	A71243	probable maltose/m
28	65.5	12.1	1321	2	T10929	3C3.20c protein -
29	65.5	12.1	1687	2	T30244	phosphodiesterase
30	65.5	12.1	1706	2	T30175	exoribonuclease, v
31	65.5	12.1	1719	2	T30174	exoribonuclease, v
32	65	12.0	209	2	S73690	MG331 homolog P01_
33	65	12.0	255	2	B64175	repressor modE hom
34	65	12.0	312	2	G71146	hypothetical prote
35	65	12.0	842	2	E96641	hypothetical prote
36	65	12.0	2233	1	ZLNZP3	genome polyprotein
37	64.5	11.9	339	2	E90436	coA-ligase / coenz
38	64.5	11.9	403	2	B89808	hypothetical prote
39	64.5	11.9	454	1	XNRTY	tyrosine transamin
40	64.5	11.9	518	2	H85440	cytochrome P450-li
41	64.5	11.9	882	1	WMBEU5	gene UL5 protein -
42	64	11.8	190	2	T20725	hypothetical prote
43	64	11.8	342	1	C69395	H+-transporting AT
44	64	11.8	446	2	G81326	probable DNA repai
45	64	11.8	627	2	H96951	fusion, PTS system

ALIGNMENTS

RESULT 1
S17449
probable ligand-binding protein RYD5 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C;Accession: S17449
R;Dear, T.N.; Boehm, T.; Keverne, E.B.; Rabbitts, T.H.
EMBO J. 10, 2813-2819, 1991
A;Title: Novel genes for potential ligand-binding proteins in subregions of the olfac
A;Reference number: S17447; MUID:92007724; PMID:1915264
A;Accession: S17449
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-94 <DEA>
A;Cross-references: EMBL:X60661; NID:g57735; PIDN:CAA43068.1; PID:g57736

Query Match 64.5%; Score 349.5; DB 2; Length 94;
Best Local Similarity 75.3%; Pred. No. 7.3e-30;
Matches 70; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

```
Qy       1 MKGSRALLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAK 60
           |||| | |||||| ||| :   ||::|||:|||||||||||||||| |||||||: ||
Db       1 MKGSSA-LLVALTVLCICGLTRAEDDNEFFMEFLQTLLVGTPEELYEGPLGKYNVNDMAK 59

Qy      61 AAMTELKSCIDGLQPMHKAELVKLLVQVLGSQD 93
           ||:|||||||| |||:|| :|||||||| :|: .
Db      60 AALTELKSCIDELQPVHKEQLVKLLVQVLDAQE 92
```

RESULT 2
A36581
polychlorinated biphenyl-binding protein precursor - rat
N;Alternate names: Clara cell 10K secretory protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 20-Aug-1999
C;Accession: A36581; S10185; S21676
R;Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gillner, M.; Gusta
J. Biol. Chem. 265, 12690-12693, 1990
A;Title: Cloning, structure, and expression of a rat binding protein for polychlorina
A;Reference number: A36581; MUID:90324266; PMID:2115524
A;Accession: A36581
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-96 <NOR>
A;Cross-references: GB:J05536; NID:g206039; PIDN:AAA41817.1; PID:g206040
R;Hagen, G.; Wolf, M.; Katyal, S.L.; Singh, G.; Beato, M.; Suske, G.
Nucleic Acids Res. 18, 2939-2946, 1990
A;Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region
A;Reference number: S10185; MUID:90272398; PMID:2349092
A;Accession: S10185
A;Status: translation not shown
A;Molecule type: DNA

```
> O <
O| |O  IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file 768826-47_x.res made by spaula on Wed 14 May 103 10:07:16-PDT.


Query sequence being compared:US-09-768-826-47 (1-105)
Number of sequences searched:                  12
Number of scores above cutoff:                 12

      Results of the initial comparison of US-09-768-826-47 (1-105) with:
        File : r55778.seq
        File : x60661.seq

   100-
      -
N     -
U   50-
M     -
B     -
E     -
R     -
      -
O     -
F   10-
      -
S     -
E    5-   **
Q     -
U     -
E     -
N     -
C     -
E     -  *  *                                                        *
S     0----------------------------------------------------------------------
      ||     |     |   | |     |     |     |     |   |   |     |
SCORE 0|     7     14  | 21    28    36    43    50  |  57     64
STDEV  0               1           2           3

                          PARAMETERS

Similarity matrix       Unitary     K-tuple                    2
Translation Frame             6
Mismatch penalty              1     Joining penalty           20
Gap penalty                1.00     Window size              105
Gap size penalty           0.05
Cutoff score                  0
Randomization group           0

                       SEARCH STATISTICS

Scores:                   Mean       Median     Standard Deviation
                             9            5          17.19

Times:                     CPU                   Total Elapsed
                     00:00:00.00                 00:00:00.00

Number of residues:                      1690
Number of sequences searched:              12
Number of scores above cutoff:             12


The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.
```

The list of best scores is:

	Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
		**** 3 standard deviations above mean ****					
1.	x60661	TOIG of: x60661 check: 2527	145	64	70	3.20	1
		**** 0 standard deviation from mean ****					
2.	x60661	TOIG of: x60661 check: 2527	144	7	21	-0.12	4
3.	r55778	TOIG of: r55778 check: 2244	137	5	15	-0.23	6
4.	x60661	TOIG of: x60661 check: 2527	143	5	22	-0.23	6
5.	x60661	TOIG of: x60661 check: 2527	143	5	18	-0.23	5
6.	x60661	TOIG of: x60661 check: 2527	144	5	19	-0.23	3
7.	r55778	TOIG of: r55778 check: 2244	137	4	12	-0.29	5
8.	r55778	TOIG of: r55778 check: 2244	138	4	13	-0.29	2
9.	r55778	TOIG of: r55778 check: 2244	138	4	16	-0.29	3
10.	r55778	TOIG of: r55778 check: 2244	138	4	10	-0.29	4
11.	r55778	TOIG of: r55778 check: 2244	139	4	14	-0.29	1
12.	x60661	TOIG of: x60661 check: 2527	144	3	16	-0.35	2

1. US-09-768-826-47 (1-105)
   x60661 TOIG of: x60661 check: 2527 from: 1 to: 435

Initial Score = 64 Optimized Score = 70 Significance = 3.20
Residue Identity = 67% Matches = 71 Mismatches = 33
Gaps = 1 Conservative Substitutions = 0
Translation Frame= 1

```
        X       10        20        30        40        50        60
        MKGSRALLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMTELKS
        |||| | ||||||  |||     ||  |||| ||||||||||||||||| |||||||  |||| |||||
    VVTVMKGSSA-LLVALTVLCICGLTRAEDDNEFFMEFLQTLLVGTPEELYEGPLGKYNVNDMAKAALTELKS
        X   10        20        30.       40        50        60        70

    70        80        90        100    X
    CIDGLQPMHKAELVKLLVQVLGSQDGAGTDYKDDDDK
    ||| ||| ||  ||||||||  |       |
    CIDELQPVHKEQLVKLLVQVLDAQEDTXAAPKHGLTSASNRTSRTRGHGCAKTSASSPVNSPSNCXINKVXC
           80        90        100       110       120       130       140
```

SP

2. US-09-768-826-47 (1-105)
   x60661 TOIG of: x60661 check: 2527 from: 1 to: 435

Initial Score = 7 Optimized Score = 21 Significance = -0.12
Residue Identity = 23% Matches = 26 Mismatches = 70
Gaps = 16 Conservative Substitutions = 0
Translation Frame= 4

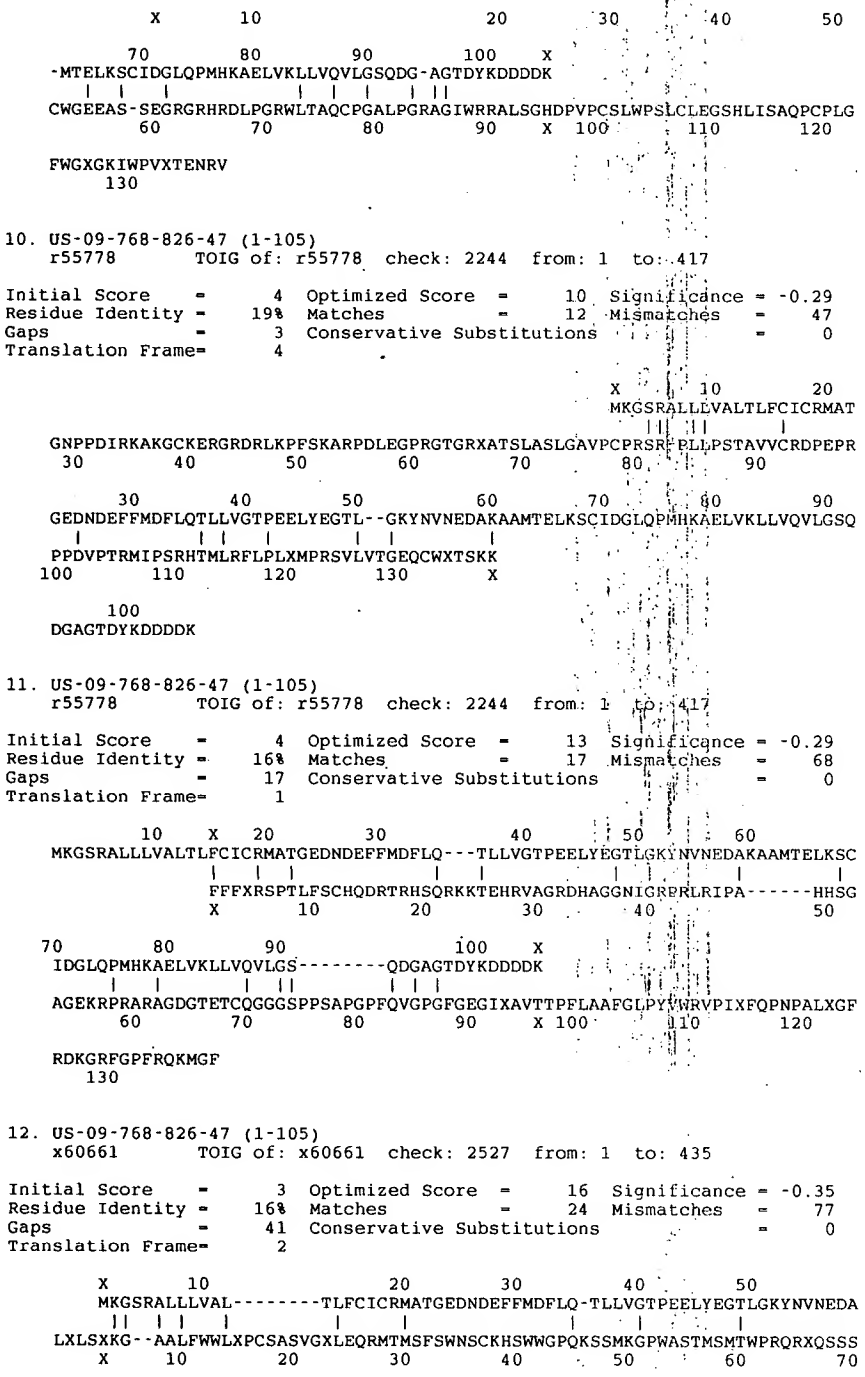
```
                                  X       10              20        30
                                  MKGSRALLLVALTLFCIC-----RMATGEDNDEFFMDFLQTL
                                    |  |    |   | |                |  ||
    WTALDFIDLTVRGXIYWXRSTGLSTSMATCTAGPIARRGQAMFWGCLCVLLSIXHLHQQLDQLLLMHRLQFI
            10        20        30        40        50        60        70

     40        50        60        70        80        90        100    X
    LVGTPEELYEGTLGKYNVNEDAKAAMTEL--KSCIDGLQPMHKAELVKLLVQVLGSQDGAGTDYKDDDDK
      |   ||    ||        |    ||        ||  |     || |    ||    ||
    NAG--LELCQRCLGHVIDIVLAQGPFIELFWGPHQECLQEFH----EKLIVILCSSQ---PTDAEHGXSHQK
            80        90        100       110             120           130  X

    SAAPFHDSH
      140
```

3. US-09-768-826-47 (1-105)
   r55778 TOIG of: r55778 check: 2244 from: 1 to: 417

Initial Score = 5 Optimized Score = 15 Significance = -0.23

```
Residue Identity -    21%  Matches          -     24  Mismatches   -    56
Gaps             -     32  Conservative Substitutions               -     0
Translation Frame-      6

          10  X     20        30        40        50        60
    MKGSRALLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMT---ELKS-
                |  |      ||        |   |   |   |   |   |       |    |||
                TRF---SVETG----QIFPYPXNPKGQGWAEIKWEPSRHK----EGQRLQGTGSXPLKSL
                X              10        20        30             40

        70        80             90             100    X
    ----CIDGLQPMHKAELVKLL-----VQVLGS------QDGAGTDYKDDDDK
            |  | | |  |  |          | |       |   |
    LQIPARPGRAPGHWA--VSHLPGNSRCRPLPSLEASSPQHRCGVPGSXAXASRCSHPHDPFPPHDAPFSSVV
  50        60          70        80        90       100       110

    NAAFCPGDRRTMLVNVKK
 120       130


4. US-09-768-826-47 (1-105)
   x60661       TOIG of: x60661  check: 2527  from: 1  to: 435

Initial Score    -      5  Optimized Score  -     22  Significance = -0.23
Residue Identity -     23%  Matches          -     26  Mismatches   -    78
Gaps             -      8  Conservative Substitutions               -     0
Translation Frame-      6

          X       10          20        30        40        50        60
          MKGSRALLLV-ALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAA--
                ||  |||       |         |         ||   | |       |   | ||
    DCIRLYXFNSSRVNLLVKKHWSXHIHGHVYGWSYCSQRSSHVLGLL-MCPLEHLALAPAAXPVAPYAPAAVH
          10        20        30        40         50        60        70

        70        80        90        100    X
    --MTELKSCIDGLQPMHKAELVKLLVQVLGSQDG--AGTDYKDDDDK
       |   |       |       |   |||  |  ||   |
    QCRTXALSALPWPCHXHCTCPGALHRALLGSPPGVFAGIPXKTHCHPLLXSAHRCRARLEPPEERCSLSXQS
          80        90       100       110       120       130       140


5. US-09-768-826-47 (1-105)
   x60661       TOIG of: x60661  check: 2527  from: 1  to: 435

Initial Score    -      5  Optimized Score  -     18  Significance = -0.23
Residue Identity -     19%  Matches          -     22  Mismatches   -    76
Gaps             -     15  Conservative Substitutions               -     0
Translation Frame-      5

        X 10        20        30        40        50        60
    MKGSRALLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMT-ELKSC--
          |   |         |||             ||    |       |   |   |   |  | ||
          GLHXTLLIXQFEGEFTGEEALVLAHPWPRVRLVLLLAE-VKPCFGAAYVSSXASSTCTSSLTSCSL
          X       10        20        30        40        50        60

       70        80             90        100    X
    ------IDGLQPMHKAELVKLLVQVL-----GSQDGAGTDYKDDDDK
                  | |   |   |       |  |  |
    CTGCSSSMQDLSSVSAALAMSLTLYLPRGPSXSSSGVPTRSVCRNSMKNSLSSSALVSPQMQSTVRATRRAL
       70        80        90       100       110 X     120       130

    LPFMTV
    140


6. US-09-768-826-47 (1-105)
   x60661       TOIG of: x60661  check: 2527  from: 1  to: 435

Initial Score    -      5  Optimized Score  -     19  Significance = -0.23
Residue Identity -     21%  Matches          -     28  Mismatches   -    62
Gaps             -     39  Conservative Substitutions               -     0
Translation Frame-      3
```

```
            X         10             20                          30
            MKGS-RALLLVA------LTLFCICRMAT-GED---------------NDEFFMDFLQTLL
              || ||| | |          |       | | |                        |  | |
    CDCHEREQRSSGGSNRALHLWADXSRGXQXVFHGIPANTPGGDPRRALXRAPGQVQCQXHGQGSADRAQVLH
          10        20        30        40        50        60        70

    40        50        60        70        80        90       100    X
    VGTPEELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGSQDGAGTDYKDDDDK
      |       |  |        | |       |  |      |  |  |  |      |   |      |
    XXT----AAGAXGATGQAAGASA------RCSRGHISSPKTWL-DLCEQ----XDQPYTWPWMCXDQCFFTS
              80        90             100       110           120    X

    KFTLELLNQXSLMQS
130       140


7. US-09-768-826-47 (1-105)
   r55778       TOIG of: r55778  check: 2244  from: 1  to: 417

Initial Score    -      4  Optimized Score  -     12  Significance = -0.29
Residue Identity -     15%  Matches          -     18  Mismatches   -    81
Gaps             -     21  Conservative Substitutions               -     0
Translation Frame-      5

                    X       10        20        30        40        50
                    MKGSRALLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGKYN
                            |         |         |                 |  |  |
    NPIFCRNGPNLPLSLKPXRAGLGXNQMGTLQTXGRPKAARNGVVTAXSPSPNPGPTWKGPGAL-GGEPPPWQ
          10        20        30        40        50        60        70

       60            70            80                    90       100
    VNEDAKAAMTELKS-----CIDGL---QPM-HKAELVKL-----------LVQVLGSQDGAGTDYKDDDDK
    |       |    | |      |   |    ||   |                |        |   |
    VSVPSPALARGLFSPAPLWCAGILSLGLPMFPPAXSLPATRCSVFFRCECRVLSWXQENNVGERQK
          80        90       100       110       120       130       X


8. US-09-768-826-47 (1-105)
   r55778       TOIG of: r55778  check: 2244  from: 1  to: 417

Initial Score    -      4  Optimized Score  -     13  Significance = -0.29
Residue Identity -     27%  Matches          -     19  Mismatches   -    28
Gaps             -     22  Conservative Substitutions               -     0
Translation Frame-      2

                                                    X       10        20
                                                    MKGSRALLLVALTLFCICRMAT
                                                      | ||   |  |  |     |
    LGSGSRHTTAVLGRRGLERGQGTAPRIAREVAHRPVPRGPSRSGRGLEKGFKRSRPRSLQPL----AFLMSG
          50        60        70        80        90 X     100

        30        40        50        60        70        80        90
    GEDNDEFFMDFLQTLLVGTPEELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGSQDG
    |       |  ||  |         |     ||     ||   |    ||
    G-----FPFDFSPT--------LPFRVLG---IRED--LARLDRKSG
  110                     120       130       X

      100
    AGTDYKDDDDK


9. US-09-768-826-47 (1-105)
   r55778       TOIG of: r55778  check: 2244  from: 1  to: 417

Initial Score    -      4  Optimized Score  -     15  Significance = -0.29
Residue Identity -     21%  Matches          -     23  Mismatches   -    60
Gaps             -     25  Conservative Substitutions               -     0
Translation Frame-      3

          10        20        30        40        50        60
    MKGSRALLLVALTLFCICRMATGEDNDEFFMDFLQTLLVGTPEELYEGTLGK-YNVNEDAKA-A--------
              ||   |       |  | |  |            | ||       ||      |   | |
           FFLTFTNIVLLSPG-QNAAF----------TTEENGASCGGKGSCGWEHREAXAQDPGTPQR
```


X 10 20 30 40 50
70 80 90 100 X
-MTELKSCIDGLQPMHKAELVKLLVQVLGSQDG-AGTDYKDDDDK
CWGEEAS-SEGRGRHRDLPGRWLTAQCPGALPGRAGIWRRALSGHDPVPCSLWPSLCLEGSHLISAQPCPLG
60 70 80 90 X 100 110 120
FWGXGKIWPVXTENRV
130
10. US-09-768-826-47 (1-105)
r55778 TOIG of: r55778 check: 2244 from: 1 to: 417
Initial Score = 4 Optimized Score = 10 Significance = -0.29
Residue Identity = 19% Matches = 12 Mismatches = 47
Gaps = 3 Conservative Substitutions = 0
Translation Frame= 4
X 10 20
MKGSRALLLVALTLFCICRMAT
GNPPDIRKAKGCKERGRDRLKPFSKARPDLEGPRGTGRXATSLASLGAVPCPRSRFPLLPSTAVVCRDPEPR
30 40 50 60 70 80 90
30 40 50 60 70 80 90
GEDNDEFFMDFLQTLLVGTPEELYEGTL--GKYNVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGSQ
PPDVPTRMIPSRHTMLRFLPLXMPRSVLVTGEQCWXTSKK
100 110 120 130 X
100
DGAGTDYKDDDDK
11. US-09-768-826-47 (1-105)
r55778 TOIG of: r55778 check: 2244 from: 1 to: 417
Initial Score = 4 Optimized Score = 13 Significance = -0.29
Residue Identity = 16% Matches = 17 Mismatches = 68
Gaps = 17 Conservative Substitutions = 0
Translation Frame= 1
10 X 20 30 40 50 60
MKGSRALLLVALTLFCICRMATGEDNDEFFMDFLQ---TLLVGTPEELYEGTLGKYNVNEDAKAAMTELKSC
FFFXRSPTLFSCHQDRTRHSQRKKTEHRVAGRDHAGGNIGRPRLRIPA------HHSG
X 10 20 30 40 50
70 80 90 100 X
IDGLQPMHKAELVKLLVQVLGS--------QDGAGTDYKDDDDK
AGEKRPRARAGDGTETCQGGGSPPSAPGPFQVGPGFGEGIXAVTTPFLAAFGLPYTWRVPIXFQPNPALXGF
60 70 80 90 X 100 110 120
RDKGRFGPFRQKMGF
130
12. US-09-768-826-47 (1-105)
x60661 TOIG of: x60661 check: 2527 from: 1 to: 435
Initial Score = 3 Optimized Score = 16 Significance = -0.35
Residue Identity = 16% Matches = 24 Mismatches = 77
Gaps = 41 Conservative Substitutions = 0
Translation Frame= 2
X 10 20 30 40 50
MKGSRALLLVAL--------TLFCICRMATGEDNDEFFMDFLQ-TLLVGTPEELYEGTLGKYNVNEDA
LXLSXKG--AALFWWLXPCSASVGXLEQRMTMSFSWNSCKHSWWGPQKSSMKGPWASTMSMTWPRQRXQSSS
X 10 20 30 40 50 60 70
60 70 80 90 100
KAAM----------TELKSCIDGLQPM----------HKAELVKLLVQVLGSQDGAGTDYKDDD
PALMNCSRCIRSNWSSCWCKCXMLKRTHKQPQNMAXPLRAIGPAVHVAMDV--LRPVLLHQXIHPRTVKSIK
80 90 100 110 120 130 140
X
DK
SNAV
X